The Krein Matrix and an Interlacing Theorem

Shamuel Auyeung [*]
Department of Mathematics and Statistics
Calvin College
Grand Rapids, MI 49546

Eric Yu [†]
Department of Mathematics and Statistics
Calvin College
Grand Rapids, MI 49546

Faculty Advisor: Todd Kapitula [‡]
Department of Mathematics and Statistics
Calvin College
Grand Rapids, MI 49546

December 10, 2013

Abstract. Consider the linear general eigenvalue problem $\boldsymbol{Ay} = \lambda \boldsymbol{By}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are both invertible and Hermitian $N \times N$ matrices. In this paper we construct a set of meromorphic functions, the Krein eigenvalues, whose zeros correspond to the real eigenvalues of the general eigenvalue problem. The Krein eigenvalues are generated by the Krein matrix, which is constructed through projections on the positive and negative eigenspaces of $\boldsymbol{B}$. The number of Krein eigenvalues depends on the number of negative eigenvalues for $\boldsymbol{B}$. These constructions not only allow for us to determine solutions to the general eigenvalue problem, but also to determine the Krein signature for each real eigenvalue. Furthermore, by applying our formulation to the simplest case of the general eigenvalue problem (where $\boldsymbol{B}$ has one negative eigenvalue), we are able to show an interlacing theorem between the eigenvalues for the general problem and the eigenvalues of $\boldsymbol{A}$.

Keywords. Krein matrix, Krein signature, linear Hermitian matrix polynomial

CONTENTS

[*]E-mail: sca6@students.calvin.edu
[†]E-mail: exy2@students.calvin.edu
[‡]E-mail: tmk5@calvin.edu

0

1. Introduction

Consider the generalized eigenvalue problem:

$$\boldsymbol{A}_{\mathrm{old}}\boldsymbol{y} = \lambda \boldsymbol{B}_{\mathrm{old}}\boldsymbol{y}, \tag{1.1}$$

where $\boldsymbol{A}_{\mathrm{old}}, \boldsymbol{B}_{\mathrm{old}} \in \mathbb{C}^{N \times N}$ are invertible Hermitian matrices. We will call the eigenvalues for this system *polynomial eigenvalues*. From an application perspective, the generalized eigenvalue problem arises when observing physical models such as coupled mass-spring systems with no damping. In this problem $\boldsymbol{A}_{\mathrm{old}}$ serves as a spring stiffness matrix, $\boldsymbol{B}_{\mathrm{old}}$ serves as the mass matrix, and λ is the square root of the natural frequency of the system [**13**]. The polynomial eigenvalues are useful in determining the stability of the system. Purely real polynomial eigenvalues imply (stable) oscillatory motion, while polynomial eigenvalues with nonzero imaginary part imply unstable exponentially growing motion. The generalized eigenvalue problem also surfaces in electrical circuit theory when using Kirchhoff's laws for coupled LC circuits. We consider another possible application near the end of this section.

The zeros of the standard characteristic polynomial, $\det(\boldsymbol{A}_{\mathrm{old}} - \lambda \boldsymbol{B}_{\mathrm{old}})$, has been a long-time standard method in determining the polynomial eigenvalues for this problem. In a recent paper published by Kollár and Miller [**11**], a new method for determining polynomial eigenvalues for linear or higher-order eigenvalue problems was defined. The eigenvalues, say $\mu_j(\lambda)$ for $j = 1, \ldots, N$, are found for the parameter-dependent matrix $\boldsymbol{A}_{\mathrm{old}} - \lambda \boldsymbol{B}_{\mathrm{old}}$, and the zeros of each of these curves correspond to polynomial eigenvalues of the original problem. Additionally, the sign of the slope of each of the $\mu_j(\lambda)$ at a zero is related to the Krein signature for the corresponding real-valued polynomial eigenvalues. If $\lambda \in \mathbb{R}$ is an algebraically and geometrically simple polynomial eigenvalue for (1.1) with corresponding eigenvector $\boldsymbol{y}$, then following Kollár and Miller [**11**, equation (1.8)] the Krein signature of the polynomial eigenvalue is given by

$$k_{\mathrm{r}}(\lambda) = -\operatorname{sign}[\boldsymbol{y}^{\mathrm{H}}\boldsymbol{B}_{\mathrm{old}}\boldsymbol{y}] \in \{-1, +1\}.$$

Here and subsequently we use the notation $\boldsymbol{C}^{\mathrm{H}} = \overline{\boldsymbol{C}^{\mathrm{T}}}$ for any matrix $\boldsymbol{C}$. The signature is said to be negative if $k_{\mathrm{r}}(\lambda) = -1$, and positive if $k_{\mathrm{r}}(\lambda) = +1$.

As discussed in Bronski et al. [**1**], Chugunova and Pelinovsky [**2**, **3**], Kapitula [**6**], Kapitula and Promislow [**7**], Kapitula et al. [**8**], Pelinovsky [**14**] and the references therein, the Krein signature of a purely real polynomial eigenvalue for the system (1.1) is a key factor in an index formula (known as the Hamiltonian-Krein index for $\star$-even matrix polynomials, see Kapitula et al. [**8**]) which counts polynomial eigenvalues based on their location in the complex plane. In particular, for the system (1.1) we have from Kapitula et al. [**8**, Theorem 4.4] (also see related results in Kollár and Miller [**11**, Theorem 5.2])

$$k_{\mathrm{r}} + k_{\mathrm{c}} + k_{\mathrm{i}} = \mathrm{n}(\boldsymbol{A}_{\mathrm{old}}). \tag{1.2}$$

Here k_{r} is the sum total of all of the purely real polynomial eigenvalues with positive Krein signature, k_{c} is the total number of complex-valued polynomial eigenvalues with nonzero real part and negative imaginary part, k_{i} is the total number of purely imaginary polynomial eigenvalues with negative imaginary part, and $\mathrm{n}(\boldsymbol{A}_{\mathrm{old}})$ denotes the number of negative eigenvalues for $\boldsymbol{A}_{\mathrm{old}}$. From a physical perspective, the Krein signature indicates a possible change in stability within the system. If two polynomial eigenvalues with the same signature collide under perturbation of the coefficient matrices, then they will pass through each other and remain purely real. On the other hand, if the polynomial eigenvalues have opposite signature, then upon collision it will generically be the case that after the collision each eigenvalue will have nonzero imaginary part. Thus, the Krein signature is useful for identifying and tracking key eigenvalues that could possibly collide and form instabilities.

In this paper we construct the Krein matrix, $\boldsymbol{K}(\lambda) \in \mathbb{C}^{k \times k}$ for $k = \mathrm{n}(\boldsymbol{B}_{\mathrm{old}})$, and from this matrix the corresponding Krein eigenvalues, $r_j(\lambda)$. The abstract formulation of the Krein matrix has been done elsewhere (e.g., see [**6**, **9**]), but here we present a systematic and concrete construction which more easily allows the interested reader to use it in a relevant problem of interest. Similar to the standard characteristic polynomial and the N eigenvalue curves $\mu_j(\lambda)$, the zeros of the k Krein eigenvalues $r_j(\lambda)$ correspond to the polynomial eigenvalues of the original linear matrix problem. Furthermore, like curves $\mu_j(\lambda)$ the sign of the derivative at zeros are related to the Krein signature of the corresponding real-valued polynomial eigenvalues.

However, unlike the Kollár-Miller approach the number of Krein eigenvalues does not depend upon the size of the matrices, and subsequently the use of the Krein eigenvalues may be a more practical method when applied to problems with large matrices.

For a particular example where the size of the matrices can be made arbitrarily large, suppose you wish to approximate the eigenvalues (denoted by χ) for the Sturm-Liouville problem

$$-\partial_x^2 y + r(x)y = \chi w(x)y, \quad y(0) = y(1) = 0,$$

where $w(x), r(x)$ are continuous. A standard approach to solving this problem is to recast this infinite-dimensional problem as a finite-dimensional one by approximating the second derivative with a matrix. This task may be accomplished in the following manner. Let N represent the number of data points, and set $h = 1/N$ to be the step size. We discretize the unit interval via $x_j = jh$ for $j = 0, \ldots, N$. We evaluate the function $y(x)$ at the data points via

$$\boldsymbol{y} := \begin{pmatrix} y(x_1) \\ y(x_2) \\ \vdots \\ y(x_{N-1}) \end{pmatrix},$$

and upon setting

$$\boldsymbol{R} := \operatorname{diag}\left(r(x_1), r(x_2), \ldots, r(x_{N-1})\right), \quad \boldsymbol{W} := \operatorname{diag}\left(w(x_1), w(x_2), \ldots, w(x_{N-1})\right),$$

we have the approximations

$$r(x)y(x) \mapsto \boldsymbol{R}\boldsymbol{y}, \quad w(x)y(x) \mapsto \boldsymbol{W}\boldsymbol{y}.$$

In order to take into account the second derivative we use the standard finite-difference approximation,

$$\partial_x^2 y(x_j) \sim \frac{y(x_{j+1}) - 2y(x_j) + y(x_{j-1})}{h^2},$$

so that

$$\partial_x^2 y(x) \mapsto \boldsymbol{D}_2 \boldsymbol{y},$$

where $\boldsymbol{D}_2$ is a symmetric and tri-diagonal matrix. The approximating eigenvalue problem is

$$\left(-\boldsymbol{D}_2 + \boldsymbol{R}\right)\boldsymbol{y} = \chi \boldsymbol{W}\boldsymbol{y}.$$

This problem can be written in our setting if we set

$$\boldsymbol{A}_{\text{old}} = \boldsymbol{W}, \quad \boldsymbol{B}_{\text{old}} = -\boldsymbol{D}_2 + \boldsymbol{R}, \quad \lambda = \frac{1}{\chi}.$$

The size of the matrices depends upon the number of discretization points.

It is often the case that for a problem of interest $\operatorname{n}\left(\boldsymbol{B}_{\text{old}}\right)$ will be a small number independent of N. For the problem outlined above it is the case that if $r(x) > 0$, so that all of the entries of the matrix $\boldsymbol{R}$ are positive, then $\boldsymbol{B}_{\text{old}}$ will be a positive-definite matrix, so that $\operatorname{n}\left(\boldsymbol{B}_{\text{old}}\right) = 0$. However, if $r(x)$ is allowed to change sign, then it is possible for there to be one, two, or even three negative eigenvalues for $\boldsymbol{B}_{\text{old}}$. For another example, in [**9**] (also see the references therein) it was the case that there were at most three negative eigenvalues for $\boldsymbol{B}_{\text{old}}$. In order to graphically determine the polynomial eigenvalues, and in particular use the graphical information to determine their Krein signature, when using the Kollár-Miller approach one must plot N curves in order to ensure that all of these polynomial eigenvalues are found. Depending on the problem, it may be the case that $N = \mathcal{O}(10^3)$, which implies that there are probably far too many curves to draw and analyze on a small number of graphs. On the other hand, when using the Krein matrix approach only a small number of curves must be determined, and this number is independent of the size of the matrices. Thus, from a visualization perspective it appears to be the case that the Krein matrix method will be easier to use.

The paper is organized as follows. In Section 2 we construct the Krein matrix and elucidate its properties. The key step, which is taken in Section 2.1, is to transform the system (1.1) into one in which the matrix $\boldsymbol{B}_{\text{old}}$

is in a canonical form. This transformation maps $\boldsymbol{A}_{\mathrm{old}} \mapsto \boldsymbol{A}$, where the new matrix satisfies the properties of being invertible with $\mathrm{n}(\boldsymbol{A}) = \mathrm{n}\,(\boldsymbol{A}_{\mathrm{old}})$. In Section 3 we prove in the case of $k = 1$ a new interlacing theorem between the eigenvalues of $\boldsymbol{A}$, a submatrix of $\boldsymbol{A}$, and the polynomial eigenvalues. This theorem can be thought of as a generalization of the classical interlacing theorems, see Horn and Johnson [**5**, Theorem 4.3.8]. We conclude with a brief discussion in Section 4.

2. The Krein matrix

2.1. *Set-up of the problem*

We start with the linear matrix problem
$$\boldsymbol{A}_{\mathrm{old}}\boldsymbol{y} = \lambda\boldsymbol{B}_{\mathrm{old}}\boldsymbol{y}, \tag{2.1}$$
where $\boldsymbol{A}_{\mathrm{old}}, \boldsymbol{B} \in \mathbb{C}^{N \times N}$ are Hermitian and invertible. If $\boldsymbol{A}_{\mathrm{old}}$ is not invertible, then $\lambda = 0$ is a polynomial eigenvalue. If $\boldsymbol{B}_{\mathrm{old}}$ is not invertible, then there would be polynomial eigenvalues at ∞: this is easily seen by setting $\chi = 1/\lambda$, considering the problem $\boldsymbol{B}_{\mathrm{old}}\boldsymbol{x} = \chi\boldsymbol{A}_{\mathrm{old}}\boldsymbol{x}$, and noting that $\chi = 0$ is a polynomial eigenvalue of this new system with associated eigenspace $\ker(\boldsymbol{B}_{\mathrm{old}})$. Thus, our assumption implies that all of the polynomial eigenvalues are finite and nonzero. Before constructing the Krein matrix for the problem, we preprocess the system so that $\boldsymbol{B}_{\mathrm{old}}$ is in a form more amenable to analysis.

Since $\boldsymbol{B}_{\mathrm{old}}$ is Hermitian, $\sigma\,(\boldsymbol{B}_{\mathrm{old}}) \subset \mathbb{R}$, where $\sigma\,(\boldsymbol{B}_{\mathrm{old}})$ denotes the collection of all eigenvalues for $\boldsymbol{B}$. We suppose that there are $1 \le k \le N - 1$ negative eigenvalues, and $N - k$ positive eigenvalues. Order the real-valued eigenvalues as
$$\mu_1 \le \mu_2 \le \cdots \le \mu_k < 0 < \mu_{k+1} \le \cdots \le \mu_N.$$
Let $\boldsymbol{v}_j$ for $j = 1, \ldots, N$ correspond to associated eigenvectors, and assume that the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal set, i.e.,
$$\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{v}_\ell = \begin{cases} 0, & j \ne \ell \\ 1, & j = \ell. \end{cases}$$

The matrix
$$\boldsymbol{Q} = (\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \cdots \ \boldsymbol{v}_n) \in \mathbb{C}^{N \times N}$$
is invertible with $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{\mathrm{H}}$. Upon defining the diagonal matrix
$$\boldsymbol{D}^{1/2} := \mathrm{diag}\left(\sqrt{|\mu_1|}, \ldots, \sqrt{|\mu_k|}, \sqrt{\mu_{k+1}}, \ldots, \sqrt{\mu_N}\right),$$
we can rewrite $\boldsymbol{B}$ in the form
$$\boldsymbol{B}_{\mathrm{old}} = \boldsymbol{C}^{\mathrm{H}} \begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix} \boldsymbol{C}, \quad \boldsymbol{C} := \boldsymbol{D}^{1/2}\boldsymbol{Q}^{\mathrm{H}}. \tag{2.2}$$

Here $\boldsymbol{I}_\ell \in \mathbb{R}^{\ell \times \ell}$ represents the identity matrix. Consequently, we can rewrite (2.1) in the form
$$\boldsymbol{C}^{\mathrm{H}} \cdot \boldsymbol{C}^{-\mathrm{H}}\boldsymbol{A}_{\mathrm{old}}\boldsymbol{C}^{-1} \cdot \boldsymbol{C}\boldsymbol{y} = \lambda\boldsymbol{C}^{\mathrm{H}} \begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix} \boldsymbol{C}\boldsymbol{y}.$$

Since $\boldsymbol{C}$, and hence $\boldsymbol{C}^{\mathrm{H}}$ are invertible, the above is equivalent to
$$\boldsymbol{C}^{-\mathrm{H}}\boldsymbol{A}_{\mathrm{old}}\boldsymbol{C}^{-1}\boldsymbol{x} = \lambda \begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix} \boldsymbol{x}, \quad \boldsymbol{x} := \boldsymbol{C}\boldsymbol{y}.$$

The matrix $\boldsymbol{C}^{-\mathrm{H}}\boldsymbol{A}_{\mathrm{old}}\boldsymbol{C}^{-1}$ is Hermitian; furthermore, by Sylvester's Law of Inertia it is the case that the number of negative/zero/positive eigenvalues for $\boldsymbol{A}_{\mathrm{old}}$ is the same as the number of negative/zero/positive eigenvalues for $\boldsymbol{C}^{-\mathrm{H}}\boldsymbol{A}_{\mathrm{old}}\boldsymbol{C}^{-1}$; in particular, $\mathrm{n}(\boldsymbol{A}) = \mathrm{n}\,(\boldsymbol{A}_{\mathrm{old}})$ [**13**, p. 568].

In conclusion, when studying the generalized eigenvalue problem (2.1) we will consider the equivalent problem

$$A x = \lambda \begin{pmatrix} -I_k & 0 \\ 0 & I_{N-k} \end{pmatrix} x, \quad A := C^{-H} A_{\mathrm{old}} C^{-1} \tag{2.3}$$

where $A \in \mathbb{C}^{N \times N}$ is Hermitian. In the case that $k = 0$ (2.3) is the regular eigenvalue problem for the matrix A; otherwise, it is a generalized eigenvalue problem.

2.2. The Krein matrix

The Krein matrix, $K(\lambda) \in \mathbb{C}^{k \times k}$, will be a matrix which has the property that under a nondegeneracy condition $\det[K(\lambda)] = 0$ if and only if λ is a polynomial eigenvalue. In all that follows we assume that we are studying the system (2.3) with

$$B = \begin{pmatrix} -I_k & 0 \\ 0 & I_{N-k} \end{pmatrix}. \tag{2.4}$$

For this matrix we have $\sigma(B) = \{-1, +1\}$. Letting $e_j \in \mathbb{R}^N$ denote the jth column of I_N, we have $\{e_1, \ldots, e_k\}$ is a basis for the negative eigenspace for B, and $\{e_{k+1}, \ldots, e_N\}$ is a basis for the positive eigenspace for B.

We start by defining the projection matrices

$$G_- := \begin{pmatrix} I_k & 0 \\ 0 & 0_{N-k} \end{pmatrix}, \quad G_+ := \begin{pmatrix} 0_k & 0 \\ 0 & I_{N-k} \end{pmatrix}.$$

The matrix G_- is a projection onto the sum of the negative eigenspace of B, and G_+ is a projection onto the positive eigenspace of B. It is clear that these projection matrices have the properties

$$G_\pm^2 = G_\pm, \quad G_- G_+ = G_+ G_- = 0_n, \quad G_- + G_+ = I_N.$$

Moreover, with respect to B we have

$$G_- B = B G_- = -G_-, \quad G_+ B = B G_+ = G_+.$$

Finally, we can write vectors as

$$x = G_- x + G_+ x = \begin{pmatrix} x_- \\ x_+ \end{pmatrix}, \tag{2.5}$$

where $x_- \in \mathbb{C}^k$ and $x_+ \in \mathbb{C}^{N-k}$.

We start our construction of the Krein matrix by using (2.5) to rewrite (2.3) as

$$A G_+ x + A G_- x = \lambda B G_- x + \lambda B G_+ x. \tag{2.6}$$

If we multiply both sides of (2.6) on the left with the projection matrix G_+, we get

$$G_+ A G_+ x + G_+ A G_- x = \lambda G_+ B G_- x + \lambda G_+ B G_+ x.$$

Upon using the facts that

$$G_+ B G_- = 0_N, \quad G_\pm^2 = G_\pm,$$

we can rewrite the above as

$$(G_+ A G_+ - \lambda G_+ B G_+) G_+ x = -G_+ A G_- \cdot G_- x. \tag{2.7}$$

We can use (2.7) to solve for x_+ in terms of x_-. The form of the projection matrices reveals that we should think of writing A in block matrix form

$$A = \begin{pmatrix} A_- & A_k^H \\ A_k & A_+ \end{pmatrix}. \tag{2.8}$$

Here $\boldsymbol{A}_- \in \mathbb{R}^{k\times k}$, $\boldsymbol{A}_+ \in \mathbb{R}^{(N-k)\times(N-k)}$ are symmetric, and $\boldsymbol{A}_k \in \mathbb{R}^{(N-k)\times k}$. It is not difficult to check that

$$\boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_- = \begin{pmatrix} \boldsymbol{A}_- & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-k} \end{pmatrix}, \quad \boldsymbol{G}_+ \boldsymbol{A}\, \boldsymbol{G}_+ = \begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_+ \end{pmatrix},$$

and

$$\boldsymbol{G}_+ \boldsymbol{A}\, \boldsymbol{G}_- = \begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{0} \\ \boldsymbol{A}_k & \boldsymbol{0}_{N-k} \end{pmatrix} = (\boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_+)^{\mathrm{T}}.$$

With this observation and the vector decomposition of (2.5) we can write (2.7) as

$$\left[\begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_+ \end{pmatrix} - \lambda \begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix} \right] \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{x}_+ \end{pmatrix} = - \begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{0} \\ \boldsymbol{A}_k & \boldsymbol{0}_{N-k} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_- \\ \boldsymbol{0} \end{pmatrix}.$$

After performing the matrix/vector multiplication the above reduces to

$$(\boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k})\, \boldsymbol{x}_+ = -\boldsymbol{A}_k \boldsymbol{x}_-. \tag{2.9}$$

The system (2.9) gives a relationship between the projection of a vector onto the positive eigenspace of $\boldsymbol{B}$ and the projection of that same vector onto the sum of the negative eigenspace and kernel of $\boldsymbol{B}$. In particular, and assuming that $\lambda \notin \sigma(\boldsymbol{A}_+)$,

$$\boldsymbol{x}_+ = -(\boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k})^{-1} \boldsymbol{A}_k \boldsymbol{x}_-. \tag{2.10}$$

Now that we have solved for $\boldsymbol{x}_+$, we must solve for $\boldsymbol{x}_-$. Multiplying both sides of (2.6) on the left by $\boldsymbol{G}_-$ gives

$$\boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_+ \boldsymbol{x} + \boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_- \boldsymbol{x} = \lambda \boldsymbol{G}_- \boldsymbol{B}\, \boldsymbol{G}_- \boldsymbol{x} + \lambda \boldsymbol{G}_- \boldsymbol{B}\, \boldsymbol{G}_+ \boldsymbol{x}.$$

Upon using the facts that

$$\boldsymbol{G}_- \boldsymbol{B}\, \boldsymbol{G}_+ = \boldsymbol{0}_N, \quad \boldsymbol{G}_\pm^2 = \boldsymbol{G}_\pm,$$

we can rewrite the above as

$$(\boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_- - \lambda \boldsymbol{G}_- \boldsymbol{B}\, \boldsymbol{G}_-)\, \boldsymbol{G}_- \boldsymbol{x} + \boldsymbol{G}_- \boldsymbol{A}\, \boldsymbol{G}_+ \cdot \boldsymbol{G}_+ \boldsymbol{x} = \boldsymbol{0}.$$

Using our previous observations regarding the actions of the projection matrices on $\boldsymbol{A}$ and $\boldsymbol{B}$ implies that the above can be written as

$$\left[\begin{pmatrix} \boldsymbol{A}_- & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-k} \end{pmatrix} - \lambda \begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-k} \end{pmatrix} \right] \begin{pmatrix} \boldsymbol{x}_- \\ \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0}_k & \boldsymbol{A}_k^{\mathrm{H}} \\ \boldsymbol{0} & \boldsymbol{0}_{N-k} \end{pmatrix} \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{x}_+ \end{pmatrix} = \boldsymbol{0}.$$

We solved for $\boldsymbol{x}_+$ in (2.10). After performing the matrix/vector multiplication the above reduces to

$$\boldsymbol{K}(\lambda)\boldsymbol{x}_- = \boldsymbol{0}, \quad \boldsymbol{K}(\lambda) := \boldsymbol{A}_- + \lambda \boldsymbol{I}_k - \boldsymbol{A}_k^{\mathrm{H}}(\boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k})^{-1} \boldsymbol{A}_k. \tag{2.11}$$

The matrix $\boldsymbol{K}(\lambda)$ is known as the *Krein matrix* [**6**].

2.3. The Krein eigenvalues

We now know that solving the generalized eigenvalue problem $(\boldsymbol{A} - \lambda \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$ is equivalent to solving $\boldsymbol{K}(\lambda)\boldsymbol{x}_- = \boldsymbol{0}$. We begin by rewriting the Krein matrix in a form that is more amenable to analysis. Since $\boldsymbol{A}_+ \in \mathbb{R}^{(N-k)\times(N-k)}$ is Hermitian, we can order the real-valued eigenvalues of this matrix, say $\lambda_1^+ \le \lambda_2^+ \le \cdots \le \lambda_{N-k}^+$. We can also assume that the associated eigenvectors, say $\boldsymbol{v}_j^+ \in \mathbb{R}^{N-k}$ for $j = 1, \dots, N-k$, form an orthonormal set. Define the rank-one projection matrices

$$\boldsymbol{G}_j^+ = \boldsymbol{v}_j^+ \left(\boldsymbol{v}_j^+ \right)^{\mathrm{H}}, \quad j = 1, \dots, N-k:$$

we have

$$\boldsymbol{G}_i^+ \boldsymbol{G}_j^+ = \begin{cases} \boldsymbol{0}_{N-k}, & i \ne j \\ \boldsymbol{G}_i^+, & i = j \end{cases}, \quad \sum_{j=1}^{N-k} \boldsymbol{G}_j^+ = \boldsymbol{I}_{N-k}.$$

The matrix $\boldsymbol{A}_+$ can be rewritten as

$$\boldsymbol{A}_+ = \sum_{j=1}^{N-k} \lambda_j^+ \boldsymbol{G}_j^+ \quad \Rightarrow \quad \boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k} = \sum_{j=1}^{N-k} (\lambda_j^+ - \lambda) \boldsymbol{G}_j^+$$

[**12**, Chapter 8]. Consequently, for $\lambda \notin \sigma(\boldsymbol{A}_+)$ the resolvent for $\boldsymbol{A}_+$ is

$$(\boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k})^{-1} = \sum_{j=1}^{N-k} \frac{1}{\lambda_j^+ - \lambda} \boldsymbol{G}_j^+. \tag{2.12}$$

Using linearity, this formulation of the resolvent gives

$$\boldsymbol{A}_k^{\mathrm{H}} (\boldsymbol{A}_+ - \lambda \boldsymbol{I}_{N-k})^{-1} \boldsymbol{A}_k = \sum_{j=1}^{N-k} \frac{1}{\lambda_j^+ - \lambda} \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{G}_j^+ \boldsymbol{A}_k.$$

The form of the projection matrix $\boldsymbol{G}_j^+$ yields

$$\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{G}_j^+ \boldsymbol{A}_k = [\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+][(\boldsymbol{v}_j^+)^{\mathrm{T}} \boldsymbol{A}_k] = \boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}}, \quad \boldsymbol{z}_j := \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+.$$

In conclusion, we can rewrite the Krein matrix as

$$\boldsymbol{K}(\lambda) = \boldsymbol{A}_- + \lambda \boldsymbol{I}_k - \sum_{j=1}^{N-k} \frac{1}{\lambda_j^+ - \lambda} \boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}}, \quad \boldsymbol{z}_j = \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+. \tag{2.13}$$

The Krein matrix has simple poles at the eigenvalues of the submatrix $\boldsymbol{A}_+$. The residue at each of the poles, λ_j^+, is the Hermitian rank-one matrix $\boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}}$.

We now wish to re-examine the system (2.11). There are two possibilities: either $\det[\boldsymbol{K}(\lambda)] = 0$, or $\boldsymbol{x}_- = \boldsymbol{0}$. Let us consider the implications of the second possibility. Going back to the form of $\boldsymbol{A}$ as written in (2.8), and using the form of $\boldsymbol{B}$ given in (2.4), we can write the original eigenvalue problem (2.6) in the form

$$\boldsymbol{A}_- \boldsymbol{x}_- + \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{x}_+ = -\lambda \boldsymbol{x}_-, \quad \boldsymbol{A}_k \boldsymbol{x}_- + \boldsymbol{A}_+ \boldsymbol{x}_+ = \lambda \boldsymbol{x}_+. \tag{2.14}$$

If $\boldsymbol{x}_- = \boldsymbol{0}$, then the system collapses to

$$\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{x}_+ = \boldsymbol{0}, \quad \boldsymbol{A}_+ \boldsymbol{x}_+ = \lambda \boldsymbol{x}_+.$$

In order for there to be a nontrivial solution, it must be the case that for some j, $\lambda = \lambda_j^+$ with $\boldsymbol{x}_+ = \boldsymbol{v}_j^+$. The equation on the left then reads $\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+ = \boldsymbol{0}$. Going back to the definition of the residue matrix at $\lambda = \lambda_j^+$ in (2.13), we then see that $\boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}} = \boldsymbol{0}_{n-k}$, i.e., the Krein matrix has a removable singularity at $\lambda = \lambda_j^+$. Conversely, suppose that the Krein matrix has a removable singularity at $\lambda = \lambda_j^+$ for some j, i.e., $\boldsymbol{z}_j = \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+ = \boldsymbol{0}$. Substituting $\boldsymbol{x}_+ = \boldsymbol{v}_j^+$ in (2.14) gives

$$\boldsymbol{A}_- \boldsymbol{x}_- = -\lambda \boldsymbol{x}_-, \quad \boldsymbol{A}_k \boldsymbol{x}_- + \boldsymbol{A}_+ \boldsymbol{v}_j^+ = \lambda \boldsymbol{v}_j^+.$$

This system clearly has a solution with $\boldsymbol{x}_- = \boldsymbol{0}$ and $\lambda = \lambda_j^+$.

Lemma 2.1. *The Krein matrix has a removable singularity at $\lambda = \lambda_j^+$, i.e., $\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+ = \boldsymbol{0}$, if and only if $\lambda_j^+ \in \sigma(\boldsymbol{A}_+)$ is a polynomial eigenvalue with associated eigenvector*

$$\boldsymbol{x} = \left(\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{v}_j^+ \end{array} \right).$$

For the sake of exposition we will henceforth assume that the Krein matrix has no removable singularities. This assumption implies that polynomial eigenvalues are found by solving $\det[\boldsymbol{K}(\lambda)] = 0$. We will further assume that the eigenvalues of $\boldsymbol{A}_+$ are strictly ordered.

Assumption 2.2. The residues of the Krein matrix, $\boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}}$ with $\boldsymbol{z}_j = \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+$, are nonzero rank-one matrices, i.e.,

$$\boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+ \neq \boldsymbol{0}, \quad j = 1, \ldots, N - k.$$

Furthermore, the real-valued poles are strictly ordered with

$$\lambda_1^+ < \lambda_2^+ < \cdots < \lambda_{N-k}^+.$$

Remark 2.3. In applications it sometimes happen that the Krein matrix has removable singularities (e.g., see Kapitula et al. [**9**]). This is not problematic, as the associated polynomial eigenvalues are generally not of interest, as they do not contribute to instabilities.

As a consequence of Assumption 2.2 and Lemma 2.1 we know that polynomial eigenvalues are found by solving $\det[\boldsymbol{K}(\lambda)] = 0$. Now, $\det[\boldsymbol{K}(\lambda)]$ is a meromorphic function with simple poles at the eigenvalues of the submatrix $\boldsymbol{A}_+$. It is a complicated function, and difficult to understand once the matrices become sufficiently large. It will be more illuminating to consider $\det[\boldsymbol{K}(\lambda)]$ via a different formulation.

It is clear that $\boldsymbol{K}(\lambda)^{\mathrm{H}} = \boldsymbol{K}(\lambda)$ for *real* λ, so that the Krein matrix is real Hermitian. This fact implies that for real λ the Krein matrix has real-valued eigenvalues, $r_j(\lambda)$, for $j = 1, \ldots, N - k$. These will hereafter be known as the *Krein eigenvalues*. The Krein eigenvalues inherit the properties of the Krein matrix; in particular, they are real meromorphic with poles at the poles of the Krein matrix [**10**, Chapter II.6]. Since the determinant of a matrix is the product of the eigenvalues,

$$\det[\boldsymbol{K}(\lambda)] = \prod_{j=1}^{k} r_j(\lambda),$$

we know that the polynomial eigenvalues are found by locating the zeros of the Krein eigenvalues.

Remark 2.4. As mentioned in Section 1, Kollár and Miller [**11**] have recently considered finding the polynomial eigenvalues using a similar idea. Namely, letting $\boldsymbol{P}_\ell(\lambda) \in \mathbb{C}^{N \times N}$ denote a ℓth-order matrix polynomial with Hermitian coefficients, they note that for real λ solving $\det[\boldsymbol{P}(\lambda)] = 0$ is equivalent to first solving

$$\boldsymbol{P}(\lambda)\boldsymbol{v}_j(\lambda) = \mu_j(\lambda)\boldsymbol{v}_j(\lambda), \quad j = 1, \ldots, N,$$

and then solving $\mu_j(\lambda) = 0$ for $j = 1, \ldots, N$.

2.4. Krein eigenvalues and the Krein signature

We now relate properties of the Krein eigenvalues to the Krein signature of a polynomial eigenvalue. For real λ we have

$$\boldsymbol{K}(\lambda)\boldsymbol{v}_j(\lambda) = r_j(\lambda)\boldsymbol{v}_j(\lambda), \quad j = 1, \ldots, k. \tag{2.15}$$

Suppose that for some $1 \leq j \leq k$ and some $\lambda_0 \in \mathbb{R}$,

$$r_j(\lambda_0) = 0 \quad \Rightarrow \quad \boldsymbol{K}(\lambda_0)\boldsymbol{v}_j(\lambda_0) = \boldsymbol{0}.$$

Recalling the vector decomposition

$$\boldsymbol{x} = \left(\begin{array}{c} \boldsymbol{x}_- \\ \boldsymbol{x}_+ \end{array} \right),$$

this implies for the eigenvector associated with the polynomial eigenvalue has $\boldsymbol{x}_- = \boldsymbol{v}_j(\lambda_0)$. Going back to (2.10) we then know

$$\boldsymbol{x}_+ = -\left(\boldsymbol{A}_+ - \lambda_0 \boldsymbol{I}_{N-k}\right)^{-1} \boldsymbol{A}_k \boldsymbol{v}_j(\lambda_0),$$

so that the eigenvector associated with the polynomial eigenvalue can be written as

$$\boldsymbol{x}_{\lambda_0} = \left(\begin{array}{c} \boldsymbol{v}_j(\lambda_0) \\ -\left(\boldsymbol{A}_+ - \lambda_0 \boldsymbol{I}_{N-k}\right)^{-1} \boldsymbol{A}_k \boldsymbol{v}_j(\lambda_0) \end{array} \right). \tag{2.16}$$

Following Gohberg et al. [**4**, Chapter 1.6] there is a Jordan chain of length at least two at the polynomial eigenvalue λ_0 if there is a solution to the set of equations

$$(\boldsymbol{A} - \lambda_0\boldsymbol{B})\boldsymbol{x}_{\lambda_0} = \boldsymbol{0}, \quad (\boldsymbol{A} - \lambda_0\boldsymbol{B})\boldsymbol{x}_1 = \boldsymbol{B}\boldsymbol{x}_{\lambda_0}.$$

Since $\boldsymbol{A} - \lambda_0\boldsymbol{B}$ is Hermitian for λ_0 real, the second equation has a solution if and only if

$$\boldsymbol{B}\boldsymbol{x}_{\lambda_0} \in \ker(\boldsymbol{A} - \lambda_0\boldsymbol{B})^\perp \quad \Rightarrow \quad \boldsymbol{x}_{\lambda_0}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}_{\lambda_0} = 0.$$

Following Kollár and Miller [**11**, equation (1.8)] we will denote

$$k_{\mathrm{r}}(\lambda_0) := -\operatorname{sign}[\boldsymbol{x}_{\lambda_0}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}_{\lambda_0}] \tag{2.17}$$

to be the *Krein signature* of the real-valued polynomial eigenvalue λ_0. If $k_{\mathrm{r}}(\lambda_0)$ is nonzero, then there is no Jordan chain.

Remark 2.5. Using the form of $\boldsymbol{B}$ yields that if we write

$$\boldsymbol{x}_{\lambda_0} = \begin{pmatrix} \boldsymbol{x}_- \\ \boldsymbol{x}_+ \end{pmatrix}; \quad \boldsymbol{x}_- \in \mathbb{C}^k, \; \boldsymbol{x}_+ \in \mathbb{C}^{N-k},$$

then

$$k_{\mathrm{r}}(\lambda_0) = -\operatorname{sign}[-\boldsymbol{x}_-^{\mathrm{H}}\boldsymbol{x}_- + \boldsymbol{x}_+^{\mathrm{H}}\boldsymbol{x}_+].$$

Remark 2.6. If λ_j^+ is a removable singularity of the Krein matrix, then upon using the characterization of the associated eigenvector as given in Lemma 2.1, i.e.,

$$\boldsymbol{x}_{\lambda_j^+} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{v}_j^+ \end{pmatrix},$$

we see from Remark 2.5 that the Krein signature of the eigenvalue is

$$k_{\mathrm{r}}(\lambda_j^+) = -\operatorname{sign}[(\boldsymbol{v}_j^+)^{\mathrm{H}}\boldsymbol{v}_j^+] = -1.$$

Thus, eigenvalues with positive Krein signature cannot appear at removable singularities of the Krein matrix.

Can the Krein signature of a polynomial eigenvalue be detected by a Krein eigenvalue? In computing the Krein signature, upon using the eigenvector as given in (2.16) and the form of the matrix $\boldsymbol{B}$ (recall (2.4)) we have

$$\boldsymbol{x}_{\lambda_0}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}_{\lambda_0} = -\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}}\boldsymbol{I}_k\boldsymbol{v}_j(\lambda_0) + [(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}\boldsymbol{A}_k\boldsymbol{v}_j(\lambda_0)]^{\mathrm{H}}(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}\boldsymbol{A}_k\boldsymbol{v}_j(\lambda_0)$$

$$= -\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}}\left[\boldsymbol{I}_k - \boldsymbol{A}_k^{\mathrm{H}}(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}\boldsymbol{A}_k\right]\boldsymbol{v}_j(\lambda_0).$$

In the second line we used the fact that $\lambda_0 \in \mathbb{R}$ implies that

$$(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-H} = (\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}.$$

From the resolvent formulation of (2.12), and the fact that the projection matrices satisfy

$$\boldsymbol{G}_i^+\boldsymbol{G}_j^+ = \begin{cases} \boldsymbol{0}_{N-k}, & i \neq j \\ \boldsymbol{G}_j^+, & i = j, \end{cases}$$

we can write

$$(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1}(\boldsymbol{A}_+ - \lambda_0\boldsymbol{I}_{N-k})^{-1} = \sum_{j=1}^{N-k}\frac{1}{(\lambda_j^+ - \lambda_0)^2}\boldsymbol{G}_j^+.$$

Thus, we have

$$\boldsymbol{x}_{\lambda_0}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}_{\lambda_0} = -\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}}\left[\boldsymbol{I}_k - \sum_{j=1}^{N-k}\frac{1}{(\lambda_j^+ - \lambda_0)^2}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{G}_j^+\boldsymbol{A}_k\right]\boldsymbol{v}_j(\lambda_0).$$

Using the definition of $\boldsymbol{z}_j$ given in (2.13) we conclude with

$$\boldsymbol{x}_{\lambda_0}^{\mathrm{H}} \boldsymbol{B} \boldsymbol{x}_{\lambda_0} = -\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}} \left[\boldsymbol{I}_k - \sum_{j=1}^{N-k} \frac{1}{(\lambda_j^+ - \lambda_0)^2} \boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}} \right] \boldsymbol{v}_j(\lambda_0). \tag{2.18}$$

The form of the Krein matrix given in (2.13) yields the observation that

$$\boldsymbol{K}'(\lambda) = \boldsymbol{I}_k - \sum_{j=1}^{N-k} \frac{1}{(\lambda_j^+ - \lambda)^2} \boldsymbol{z}_j \boldsymbol{z}_j^{\mathrm{H}}.$$

This observation being substituted into (2.18) with $\lambda = \lambda_0$ gives the relationship,

$$\boldsymbol{x}_{\lambda_0}^{\mathrm{H}} \boldsymbol{B} \boldsymbol{x}_{\lambda_0} = -\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}} \boldsymbol{K}'(\lambda_0) \boldsymbol{v}_j(\lambda_0).$$

From (2.17) we conclude with

$$k_{\mathrm{r}}(\lambda_0) = \mathrm{sign}[\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}} \boldsymbol{K}'(\lambda_0) \boldsymbol{v}_j(\lambda_0)]. \tag{2.19}$$

We finally wish to relate the derivative of a Krein eigenvalue at a zero to the Krein signature of the eigenvalue. Since the Krein matrix and Krein eigenvalues are real meromorphic we can differentiate (2.15) with respect to λ to get

$$\boldsymbol{K}'(\lambda)\boldsymbol{v}_j(\lambda) + \boldsymbol{K}(\lambda)\boldsymbol{v}_j'(\lambda) = r_j'(\lambda)\boldsymbol{v}_j(\lambda) + r_j(\lambda)\boldsymbol{v}_j'(\lambda).$$

Evaluating at $\lambda = \lambda_0$, using $r_j(\lambda_0) = 0$, and rearranging gives

$$\boldsymbol{K}(\lambda_0)\boldsymbol{v}_j'(\lambda_0) = r_j'(\lambda_0)\boldsymbol{v}_j(\lambda_0) - \boldsymbol{K}'(\lambda_0)\boldsymbol{v}_j(\lambda_0).$$

This linear system must have a solution, as $\boldsymbol{v}_j'(\lambda_0)$ exists (recall that the Krein eigenvalues and associated eigenvectors are real meromorphic, which implies the existence of derivatives of all orders away from a pole). Since $\boldsymbol{K}(\lambda_0)$ is Hermitian, this means that

$$r_j'(\lambda_0)\boldsymbol{v}_j(\lambda_0) - \boldsymbol{K}'(\lambda_0)\boldsymbol{v}_j(\lambda_0) \in \ker[\boldsymbol{K}(\lambda_0)]^{\perp},$$

i.e.,

$$\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}}[r_j'(\lambda_0)\boldsymbol{v}_j(\lambda_0) - \boldsymbol{K}'(\lambda_0)\boldsymbol{v}_j(\lambda_0)] = 0.$$

Using linearity and simplifying gives the relationship

$$r_j'(\lambda_0) = \frac{\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}} \boldsymbol{K}'(\lambda_0)\boldsymbol{v}_j(\lambda_0)}{\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}}\boldsymbol{v}_j(\lambda_0)}.$$

Since the denominator is positive,

$$\mathrm{sign}[r_j'(\lambda_0)] = \mathrm{sign}[\boldsymbol{v}_j(\lambda_0)^{\mathrm{H}} \boldsymbol{K}'(\lambda_0)\boldsymbol{v}_j(\lambda_0)]. \tag{2.20}$$

This calculation, along with the relationship (2.19), gives us the following result:

Lemma 2.7. *Under Assumption 2.2 the polynomial eigenvalues are realized as a zero of (at least) one of the Krein eigenvalues. The polynomial eigenvalue will be algebraically simple, i.e., there will be no Jordan chain, if and only if the zero of the Krein eigenvalue is simple. Furthermore, the sign of the derivative of the Krein eigenvalue is related to the Krein signature of the real-valued polynomial eigenvalue (see (2.17)) by*

$$\mathrm{sign}[r_j'(\lambda_0)] = k_{\mathrm{r}}(\lambda_0).$$

In other words, if the Krein signature is $+1$, then the graph of the Krein eigenvalue has positive slope at the zero, and if the Krein signature is -1, then the graph of the Krein eigenvalue has negative slope at the zero.

Recall from Section 1 that the notion of a Krein signature for real-valued polynomial eigenvalues allows us to count the number of real, purely imaginary, and complex (with nonzero imaginary part) polynomial eigenvalues in terms of the matrix coefficients $\boldsymbol{A}$ and $\boldsymbol{B}$. In particular, we have

$$k_{\mathrm{r}} + k_{\mathrm{c}} + k_{\mathrm{i}} = \mathrm{n}\left(\boldsymbol{A}\right), \tag{2.21}$$

where again k_{r} is the total number of real-valued polynomial eigenvalues with positive Krein signature. Using the graphical result of Lemma 2.7 with the index result of (2.21), we see that we can obtain a good idea of the spectrum for (2.3) simply by plotting all of the Krein eigenvalues for real λ, observing the slopes of their intersections with the λ-axis, and counting.

Remark 2.8. Recall that in Section 2.1 we showed that by Sylvester's Law of Inertia $\mathrm{n}(\boldsymbol{A}) = \mathrm{n}(\boldsymbol{A}_{\mathrm{old}})$. Furthermore, we have

$$\boldsymbol{x}^{\mathrm{H}}\begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix}\boldsymbol{x} = (\boldsymbol{C}\boldsymbol{y})^{\mathrm{H}}\begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix}(\boldsymbol{C}\boldsymbol{y})$$
$$= \boldsymbol{y}^{\mathrm{H}}\left[\boldsymbol{C}^{\mathrm{H}}\begin{pmatrix} -\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-k} \end{pmatrix}\boldsymbol{C}\right]\boldsymbol{y}$$
$$= \boldsymbol{y}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{y}$$

(going from the second to third line we are recalling (2.2)). Consequently, the Krein signature is invariant under the transformations leading from the original problem (2.1) to the transformed problem (2.3). While (2.21) is stated in terms of the transformed problem (2.3), it is the same for the original problem.

2.5. Plots of the Krein eigenvalues

We now present representative plots of the Krein eigenvalues for various values of k. In all of the examples we will assume that $N = 8$.

We first assume $k = 1$, so that the Krein matrix is the Krein eigenvalue. In Figure 1 we see two representative plots for a given matrix $\boldsymbol{A}$ which satisfies $\mathrm{n}(\boldsymbol{A}) = 1$. From (2.21) the eigenvalues satisfy the distribution

$$k_{\mathrm{r}} + k_{\mathrm{c}} + k_{\mathrm{i}} = 1.$$

The top panel is a plot of the polynomial eigenvalues, and the bottom is a plot of the Krein eigenvalue. We see that for this problem $k_{\mathrm{r}} = 1$ with $k_{\mathrm{c}} = k_{\mathrm{i}} = 0$.

Now assume $k = 2$, so that there are two Krein eigenvalues to be plotted. In Figure 2 we see two representative plots for a given matrix $\boldsymbol{A}$ which satisfies $\mathrm{n}(\boldsymbol{A}) = 2$. From (2.21) the eigenvalues satisfy the distribution

$$k_{\mathrm{r}} + k_{\mathrm{c}} + k_{\mathrm{i}} = 2.$$

The top panel is a plot of the polynomial eigenvalues, and the bottom is a plot of the Krein eigenvalue. We see that for this problem $k_{\mathrm{r}} = k_{\mathrm{c}} = 1$ with $k_{\mathrm{i}} = 0$. Note that in this plot of the Krein eigenvalues there is the interesting feature that each pole of the Krein matrix is a pole for only one of the Krein eigenvalues. This observation is a consequence of the fact that the residue of the Krein matrix at each pole is for this problem a rank one matrix. Consequently, one Krein eigenvalue will have a removable singularity at the pole of the Krein matrix. In general, it will be the case that for any k only one of the Krein eigenvalues will have a pole where the Krein matrix has the pole, and the remaining $k - 1$ Krein eigenvalues will treat the pole as a removable singularity.

3. AN INTERLACING THEOREM

Our goal in this section is to show a relationship between the polynomial eigenvalues, the eigenvalues of $\boldsymbol{A}$, and the eigenvalues of the submatrix $\boldsymbol{A}_+$. In addition to Assumption 2.2, we will in this section assume that $k = 1$. Going back to the decomposition of the matrix $\boldsymbol{A}$ given in (2.8), we will have $\boldsymbol{A}_- = \boldsymbol{A}_{11}$ (that entry
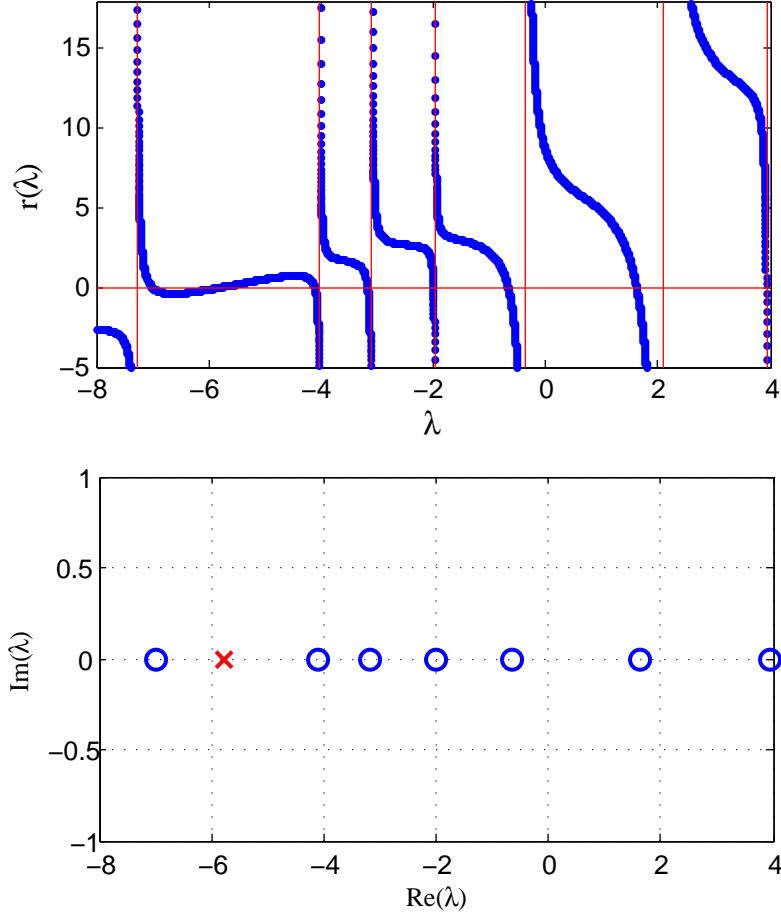
Figure 1: (color online) A numerically generated plot of the Krein eigenvalue (top panel) and the polynomial eigenvalues (bottom panel) for the case that $k = 1$ and $N = 8$. The vertical asymptotes are denoted by thin (red) lines, and are located at the eigenvalues for the submatrix $\boldsymbol{A}_+$. In the bottom panel the real-valued polynomial eigenvalues with negative signature are denoted by (blue) circles, and those with positive signature by (red) crosses. All of the polynomial eigenvalues are real-valued, and all but one have negative Krein signature.

in the matrix $\boldsymbol{A}$), $\boldsymbol{A}_+ \in \mathbb{C}^{(N-1)\times(N-1)}$, and $\boldsymbol{A}_k \in \mathbb{C}^{N-1}$. The Krein matrix, $\boldsymbol{K}(\lambda)$, becomes a meromorphic function, and is consequently the Krein eigenvalue $r(\lambda)$, which is given by

$$r(\lambda) = A_{11} + \lambda - \sum_{j=1}^{N-1} \frac{|z_j|^2}{\lambda_j^+ - \lambda}, \quad z_j = \boldsymbol{A}_k^{\mathrm{H}} \boldsymbol{v}_j^+ \in \mathbb{C}\backslash\{0\}. \tag{3.1}$$

Here A_{11} refers to that entry in the matrix $\boldsymbol{A}$, and $z_j \neq 0$ follows from Assumption 2.2. By Assumption 2.2 all of the polynomial eigenvalues are realized as zeros of $r(\lambda)$.

Our interlacing theorem requires that we locate the eigenvalues of $\boldsymbol{A}$ relative to those of the submatrix $\boldsymbol{A}_+$. This task will be accomplished by constructing a meromorphic function, say $s(\lambda)$, which has a form similar to $r(\lambda)$, but whose zeros correspond to the eigenvalues for the Hermitian matrix $\boldsymbol{A}$. We start with the eigenvalue problem (2.3), except now the matrix $\boldsymbol{B}$ is replace by $\boldsymbol{I}_n$ (i.e., $k = 0$). By using the projection matrices

$$\boldsymbol{G}_- := \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-1} \end{pmatrix}, \quad \boldsymbol{G}_+ := \begin{pmatrix} 0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{N-1} \end{pmatrix},$$

and following the reasoning in Section 2.2, we see that there is no change in the vector expression for $\boldsymbol{x}_+$
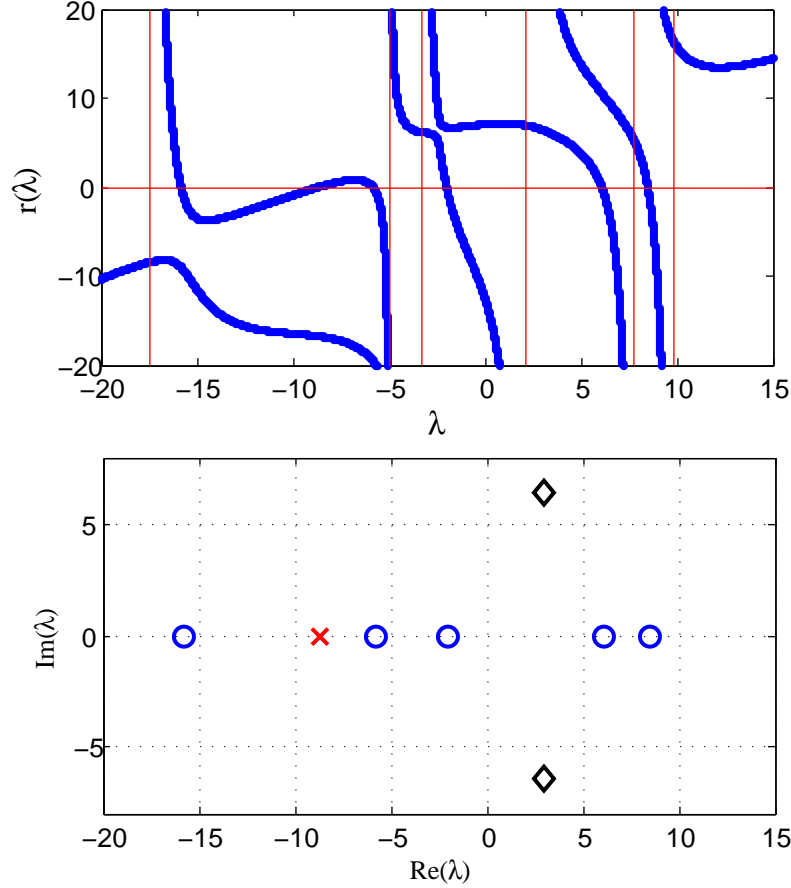
Figure 2: (color online) A numerically generated plot of the Krein eigenvalue (top panel) and the polynomial eigenvalues (bottom panel) for the case that $k = 2$ and $N = 8$. The vertical asymptotes are denoted by thin (red) lines, and are located at the eigenvalues for the submatrix $\boldsymbol{A}_+$. In the bottom panel the real-valued polynomial eigenvalues with negative signature are denoted by (blue) circles, and those with positive signature by (red) crosses. There are six real-valued polynomial eigenvalues, only one of which has positive Krein signature. The remaining two polynomial eigenvalues are complex-valued with nonzero real and imaginary part.

given in (2.9). Indeed, the only difference is that the line above (2.11) for $\boldsymbol{B} = \boldsymbol{I}_N$ becomes

$$\left[\begin{pmatrix} A_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-1} \end{pmatrix} - \lambda \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{N-1} \end{pmatrix}\right]\begin{pmatrix} x_- \\ \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} 0 & \boldsymbol{A}_k^{\mathrm{T}} \\ \boldsymbol{0} & \boldsymbol{0}_{N-1} \end{pmatrix}\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{x}_+ \end{pmatrix} = \boldsymbol{0}.$$

Substitution of the expression for $\boldsymbol{x}_+$ into the above, and using the resolvent expression of (2.12) along with the resultant simplifications, yields

$$s(\lambda)x_- = 0, \quad s(\lambda) := A_{11} - \lambda - \sum_{j=1}^{N-1} \frac{|z_j|^2}{\lambda_j^+ - \lambda}, \quad z_j = \boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{v}_j^+. \tag{3.2}$$

By Assumption 2.2 it must be true that $x_- \neq 0$ for any eigenvector, so that in conclusion an eigenvalue for $\boldsymbol{A}$ is found if and only if $s(\lambda) = 0$.

Comparing the functions $r(\lambda)$ and $s(\lambda)$, we see that the meromorphic part of both is the common function

$$m(\lambda) := \sum_{j=1}^{N-1} \frac{|z_j|^2}{\lambda_j^+ - \lambda}, \quad z_j \neq 0. \tag{3.3}$$

By Assumption 2.2 the simple poles of $m(\lambda)$ are ordered as

$$\lambda_1^+ < \lambda_2^+ < \cdots < \lambda_{N-1}^+,$$

and the residue of each pole, $|z_j|^2$, is positive. For real λ the function $m(\lambda)$ has the following properties:

(a) $\displaystyle\lim_{\lambda\to\pm\infty} m(\lambda) = 0$

(b) $\displaystyle m'(\lambda) = \sum_{j=1}^{N-1} \frac{|z_j|^2}{(\lambda_j^+ - \lambda)^2} > 0.$

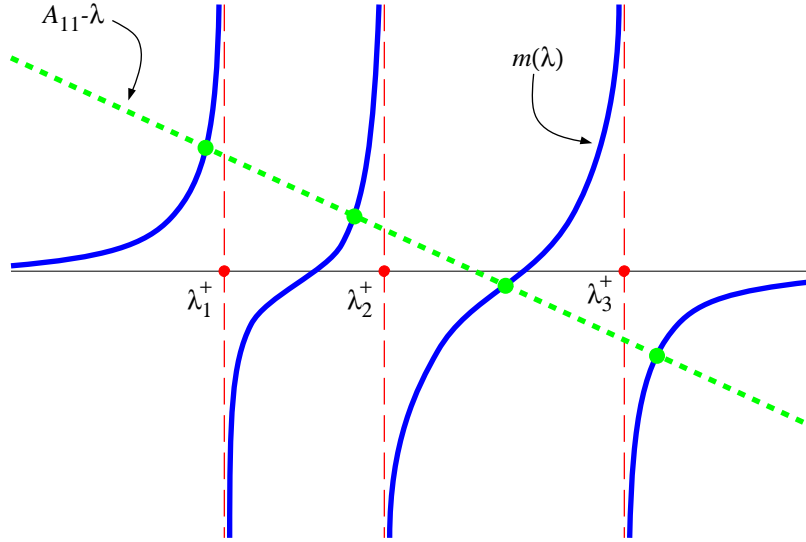Property (b) implies that $m(\lambda)$ is strictly increasing.

Figure 3: (color online) A cartoon showing showing the plots of $m(\lambda)$ (solid (blue) curve) and $A_{11} - \lambda$ (dashed (green) curve) for real-valued λ when $N = 4$. The intersection points correspond to zeros of $s(\lambda)$, which are the eigenvalues of $\boldsymbol{A}$.

Denote the zeros of $s(\lambda)$ - the eigenvalues of $\boldsymbol{A}$ - as λ_j^S for $j = 1, \ldots, N$. These zeros are realized as the intersection of the curves $A_{11} - \lambda$ and $m(\lambda)$, i.e.,

$$A_{11} - \lambda = m(\lambda).$$

The curve on the left side of the above equation is a straight line with negative slope. It is clear that this line will intersect $m(\lambda)$ at precisely N points; furthermore, there will be precisely one intersection between each of the poles, one intersection to the left of λ_1^+, and one intersection to the right of λ_{n-1}^+ (see Figure 3). In conclusion, upon recalling that

(a) the poles of $m(\lambda)$ are the eigenvalues of the submatrix $\boldsymbol{A}_+$

(b) the zeros of $s(\lambda)$ are the eigenvalues of $\boldsymbol{A}$,

we have the ordering

$$\lambda_1^S < \lambda_1^+ < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S. \tag{3.4}$$

This is a classical result (see Horn and Johnson [5, Theorem 4.3.8]): the point to be made here is that the assumption $\boldsymbol{A}_k^{\mathrm{T}} \boldsymbol{v}_j^+ \neq 0$ for all j implies that the ordering is strict. If this assumption is removed, then some of the eigenvalues for $\boldsymbol{A}_+$ will coincide with those of $\boldsymbol{A}$.
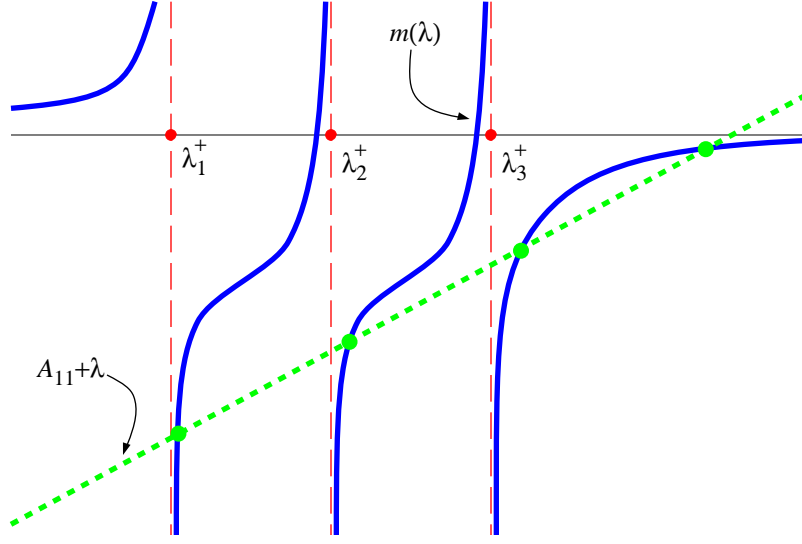
Figure 4: (color online) A cartoon showing showing the plots of $m(\lambda)$ (solid (blue) curve) and $A_{11} + \lambda$ (dashed (green) curve) for real-valued λ when $N = 4$. The intersection points correspond to zeros of $r(\lambda)$, which are the polynomial eigenvalues.

Now consider the location of the polynomial eigenvalues. Denote the zeros of $r(\lambda)$ - the polynomial eigenvalues - as λ_j^K for $j = 1, \ldots, N$. These zeros are realized as the intersection of the curves $A_{11} + \lambda$ and $m(\lambda)$, i.e.,

$$A_{11} + \lambda = m(\lambda).$$

The curve on the left side of the above equation is a straight line with positive slope. This line will always intersect $m(\lambda)$ in at least one point between each of the poles. If $-A_{11} > 0$ is sufficiently large, i.e., A_{11} is negative with large enough absolute value, then the line will intersect $m(\lambda)$ in two distinct points to the right of λ_{N-1}^+, and there will be only one intersection point between each of the poles (see Figure 4)). This assumption provides the ordering

$$\lambda_1^+ < \lambda_1^K < \lambda_2^+ < \lambda_2^K < \cdots \lambda_{N-2}^+ < \lambda_{N-2}^K < \lambda_{N-1}^+ < \lambda_{N-1}^K < \lambda_n^K. \tag{3.5}$$

On the other hand, the line will intersect the curve to the left of the first pole λ_1^+ in two distinct points if $A_{11} > 0$ is sufficiently large, and there will be only one intersection point between each of the poles. Thus, if $A_{11} > 0$ is sufficiently large all of the polynomial eigenvalues are real-valued, and we have the ordering

$$\lambda_1^K < \lambda_2^K < \lambda_1^+ < \lambda_3^K < \lambda_2^+ < \cdots \lambda_{N-2}^+ < \lambda_N^K < \lambda_{N-1}^+. \tag{3.6}$$

The interlacing results of (3.4)-(3.5) can be made into a result which combines the eigenvalues for all three problems. The fact that $m(\lambda)$ is a strictly increasing function implies that intersections of that curve with $A_{11} - \lambda$ will always lie to the left of the intersections of that curve with $A_{11} + \lambda$ (compare Figure 3 with Figure 4). Thus, if $A_{11} > 0$ is sufficiently large the orderings of (3.4) and (3.6) can be combined to

$$\lambda_1^K < \lambda_2^K < \lambda_1^S < \lambda_1^+ < \lambda_3^K < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_N^K < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S.$$

On the other hand, if $-A_{11} > 0$ is sufficiently large, then the orderings of (3.4) and (3.5) can be combined to

$$\lambda_1^S < \lambda_1^+ < \lambda_1^K < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_{N-2}^+ < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S < \lambda_{N-1}^K < \lambda_N^K.$$

Theorem 3.1. *Suppose that Assumption 2.2 holds with $k = 1$. The eigenvalues of $\boldsymbol{A}$, λ_j^S for $j = 1, \ldots, N$, are related to the eigenvalues of the submatrix $\boldsymbol{A}_+$ through the interlacing*

$$\lambda_1^S < \lambda_1^+ < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S.$$

If $A_{11} > 0$ is sufficiently large, then the polynomial eigenvalues, λ_j^K for $j = 1, \dots, N$, satisfy the interlacing

$$\lambda_1^K < \lambda_2^K < \lambda_1^S < \lambda_1^+ < \lambda_3^K < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_N^K < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S.$$

Finally, if $-A_{11} > 0$ is sufficiently large there is the interlacing

$$\lambda_1^S < \lambda_1^+ < \lambda_1^K < \lambda_2^S < \lambda_2^+ < \cdots < \lambda_{N-2}^K < \lambda_{N-1}^S < \lambda_{N-1}^+ < \lambda_N^S < \lambda_{N-1}^K < \lambda_N^K.$$
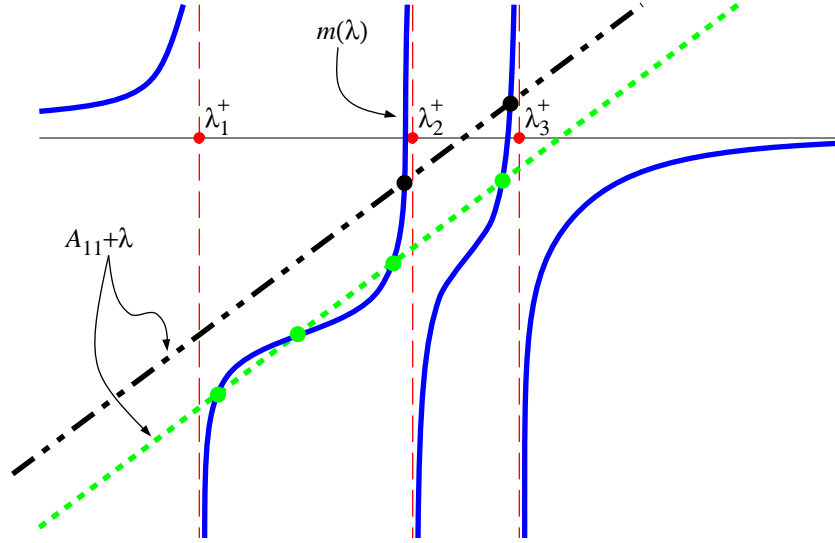

Figure 5: (color online) A cartoon showing showing the plots of $m(\lambda)$ (solid (blue) curve) and $A_{11} + \lambda$ for two different values of A_{11} when $N = 4$. The intersection points correspond to zeros of $r(\lambda)$, which are the polynomial eigenvalues. For the dashed (green) curve A_{11} is chosen so that three of the polynomial eigenvalues are between the poles λ_1^+ and λ_2^+. It is clear that if A_{11} were a little larger (dot-dashed (black) curve), then the three intersection points would become one intersection point. In this case there would be only two real-valued solutions to $r(\lambda) = 0$, which means that the other two polynomial eigenvalues must have nonzero imaginary part.

If $|A_{11}|$ is sufficiently large with $k = 1$, then all of the polynomial eigenvalues will be real-valued and have the ordering detailed by Theorem 3.1. If $|A_{11}|$ is not too large, then $A_{11} + \lambda = m(\lambda)$ will have at least one intersection point between each of the poles, but no intersection points to the left of λ_1^+ or the right of λ_{N-1}^+. There will be at least $N - 2$ real-valued polynomial eigenvalues. The remaining two will either be real-valued and located between two adjacent poles, or complex-valued with nonzero imaginary part (see Figure 5).

Remark 3.2. The parameter A_{11} is clearly most important in this analysis. In terms of the original matrices, if we recall the reformulation of the problem discussed in Section 2.1, then upon using

$$\boldsymbol{A} = \boldsymbol{C}^{-\mathrm{H}} \boldsymbol{A}_{\mathrm{old}} \boldsymbol{C}^{-1}, \quad \boldsymbol{C} = \boldsymbol{D}^{1/2} \boldsymbol{Q}^{\mathrm{H}},$$

we see that

$$A_{11} = \frac{\boldsymbol{q}_1^{\mathrm{H}} \boldsymbol{A}_{\mathrm{old}} \boldsymbol{q}_1}{|\mu_1|}.$$

Here $\mu_1 < 0$ is the (only) negative eigenvalue of $\boldsymbol{B}_{\mathrm{old}}$ with associated and normalized eigenvector $\boldsymbol{q}_1$.

4. Conclusion

We consider the linear general eigenvalue problem $\boldsymbol{A}_{\text{old}}\boldsymbol{y} = \lambda\boldsymbol{B}_{\text{old}}\boldsymbol{y}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are both invertible and Hermitian $N \times N$ matrices. The eigenvalues of the general problem can also be found by determining the eigenvalues of the non-Hermitian matrix $\boldsymbol{B}_{\text{old}}^{-1}\boldsymbol{A}_{\text{old}}$. By diagonalizing $\boldsymbol{B}_{\text{old}}$, and thinking of the reformulated matrix $\boldsymbol{A}_{\text{old}}$ as a block matrix where the size of the upper left block depends on the number of negative eigenvalues of $\boldsymbol{B}_{\text{old}}$, we are able to provide an intuitive and easily implementable construction of the Krein matrix. The eigenvalues of the Krein matrix are used to find the polynomial eigenvalues, much as the zeros of the characteristic polynomial find the eigenvalues of a given matrix. By applying our formulation to the case where $\boldsymbol{B}_{\text{old}}$ has one negative eigenvalue, we are able to show an interlacing theorem between the polynomial eigenvalues and the eigenvalues of $\boldsymbol{A}_{\text{old}}$. In particular, we show that if $\left|\boldsymbol{q}_1^{\text{H}}\boldsymbol{A}_{\text{old}}\boldsymbol{q}_1/|\mu_1|\right|$ is sufficiently large, where $\mu_1 < 0$ is the only negative eigenvalue of $\boldsymbol{B}_{\text{old}}$ with associated normalized eigenvector $\boldsymbol{q}_1$, then the eigenvalues of $\boldsymbol{B}_{\text{old}}^{-1}\boldsymbol{A}_{\text{old}}$ are real-valued and interlace with those of $\boldsymbol{A}_{\text{old}}$.

The next problem to consider is the extension of the Krein matrix to the study of quadratic eigenvalue problems,

$$\left(\boldsymbol{A} + \lambda\boldsymbol{B} + \lambda^2\boldsymbol{C}\right)\boldsymbol{y} = \boldsymbol{0},$$

where each matrix is Hermitian and invertible. Such problems occur, for example, in the study of coupled mass-spring systems with damping, coupled RLC circuits, acoustics, and flows in fluid mechanics, see Tisseur and Meerbergen [15]. This topic will be the subject of a future paper.

References

[1] J. Bronski, M. Johnson, and T. Kapitula. An instability index theory for quadratic pencils and applications. preprint, 2012.

[2] M. Chugunova and D. Pelinovsky. On quadratic eigenvalue problems arising in stability of discrete vortices. *Lin. Alg. Appl.*, 431:962–973, 2009.

[3] M. Chugunova and D. Pelinovsky. Count of eigenvalues in the generalized eigenvalue problem. *J. Math. Phys.*, 51(5):052901, 2010.

[4] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*, volume 58 of *Classics in Applied Mathematics*. SIAM, 2009.

[5] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 1985.

[6] T. Kapitula. The Krein signature, Krein eigenvalues, and the Krein Oscillation Theorem. *Indiana U. Math. J.*, 59:1245–1276, 2010.

[7] T. Kapitula and K. Promislow. *Spectral and Dynamical Stability of Nonlinear Waves*. Springer-Verlag, 2013.

[8] T. Kapitula, E. Hibma, H.-P. Kim, and J. Timkovich. Instability indices for matrix polynomials. *Linear Algebra Appl.*, 439:3412–3434, 2013.

[9] T. Kapitula, P. Kevrekidis, and D. Yan. The Krein matrix: General theory and concrete applications in atomic Bose-Einstein condensates. *SIAM J. Appl. Math.*, 73(4):1368–1395, 2013.

[10] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin, 1980.

[11] R. Kollár and P. Miller. Graphical Krein signature theory and Evans-Krein functions. preprint, 2013.

[12] P. Lax. *Linear Algebra*. John Wiley & Sons, Inc., 1997.

[13] C. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.

[14] D. Pelinovsky. Inertia law for spectral stability of solitary waves in coupled nonlinear Schrödinger equations. *Proc. Royal Soc. London A*, 461:783–812, 2005.

[15] R. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.